WHAT IS CLAIMED IS:

1. A speech synthesis apparatus comprising:

distortion output means for obtaining a distortion produced upon modifying a synthesis unit on the basis of predetermined prosody information; and

unit registration means for selecting a synthesis unit to be registered in a synthesis unit inventory used in speech synthesis on the basis of the distortion output from said distortion output means.

2. The apparatus according to claim 1, wherein said distortion output means obtains the distortion on the basis of a concatenation distortion produced upon concatenating the synthesis unit to another synthesis unit, and a modification distortion produced upon modifying the synthesis unit.

3. The apparatus according to claim 1, further comprising:

text input means for inputting text data;

language analysis means for performing language analysis of the input text data; and

prosody generation means for generating the predetermined prosody information on the basis of an analysis result of said language analysis means.

4. The apparatus according to claim 2, further comprising:

Nbest determination means for obtaining Nbest sequences of a synthesis unit sequence with reference to the distortion determined based on the concatenation and modification distortions, and

wherein said unit registration means selects a synthesis unit to be registered in the synthesis unit inventory on the basis of the Nbest sequences of the synthesis unit sequence.

5. The apparatus according to claim 2, wherein said unit registration means selects a synthesis unit to be registered in the synthesis unit inventory on the basis of a weighted sum of the concatenation and modification distortions.

6. The apparatus according to claim 2, wherein said distortion output means determines the concatenation distortion using a cepstrum distance between synthesis units.

7. The apparatus according to claim 2, wherein said distortion output means determines the modification distortion using a cepstrum distance between synthesis units before and after modification.

8. The apparatus according to claim 2, wherein said distortion output means has a table that stores the modification distortion, and determines the modification distortion by looking up the table.

9. The apparatus according to claim 2, wherein said distortion output means has a table that stores the concatenation distortion, and determines the concatenation distortion by looking up the table.

10. The apparatus according to claim 1, further comprising speech synthesis means for producing synthetic speech of text data using the synthesis unit inventory.

11. A speech synthesis method comprising:

a distortion output step of obtaining a distortion produced upon modifying a synthesis unit on the basis of predetermined prosody information; and

a unit registration step of selecting a synthesis unit to be registered in a synthesis unit inventory used in speech synthesis on the basis of the distortion output from the distortion output step.

12. The method according to claim 11, wherein in said distortion output step, the distortion is obtained on the basis of a concatenation distortion produced upon

concatenating the synthesis unit to another synthesis unit, and a modification distortion produced upon modifying the synthesis unit.

13. The method according to claim 11, further comprising the steps of:

inputting text data;

performing language analysis of the input text data; and

generating the predetermined prosody information on the basis of an analysis result in the language analysis step.

14. The method according to claim 12, further comprising the step of:

obtaining Nbest sequences of a synthesis unit sequence with reference to the distortion determined based on the concatenation and modification distortions, and

wherein in said unit registration step, a synthesis unit to be registered in the synthesis unit inventory is selected on the basis of the Nbest sequences of the synthesis unit sequence.

15. The method according to claim 12, wherein in said unit registration step, synthesis unit to be registered in the synthesis unit inventory is selected on the

basis of a weighted sum of the concatenation and modification distortions.

16. The method according to claim 12, wherein in said distortion output step, the concatenation distortion is determined by using a cepstrum distance between synthesis units.

17. The method according to claim 12, wherein in said distortion output step, the distortion is obtained by quantifying the modification distortion as a cepstrum distance between synthesis units before and after modification.

18. The method according to claim 12, wherein in said distortion output step, the modification distortion is determined by looking up a table that stores the modification distortion.

19. The method according to claim 2, wherein in said distortion output step, the concatenation distortion is determined by looking up a table that stores the concatenation distortion.

20. The method according to claim 11, further comprising a speech synthesis step of producing

synthetic speech of text data using the synthesis unit inventory.

21. A computer readable storage medium storing a program that implements a method cited in claim 11.